

PdfSlantTextCorrector

利用者マニュアル

1.5.1 版(標準版)

目次

- はじめに..... 2
- 定義..... 2
- 特徴..... 2
- バージョン 1.5.1 改訂内容..... 2
- 留意点..... 3
- 利用規約 (Terms of Use)..... 3
- 利用想定環境..... 3
- Java 環境の整備..... 4
- ライブラリの配備と使用..... 4
- インストール..... 5
- アンインストール..... 5
- 開発及びテスト環境..... 6
- 配布物..... 6
- 基本操作手順..... 6
- 画面説明..... 8
- プロパティファイル説明..... 16
- プロパティファイルの更新..... 20
- エラーファイル..... 20
- Q&A..... 20
- 連絡先..... 21

はじめに

PdfSlantTextCorrector の目的はスキャン処理等における PDF ファイルコンテンツの傾きやテキストの誤字・文字化けを補正する機能を提供することです。

ページ毎にコンテンツの種類(文書形式・画像形式等)や傾きが異なる PDF ファイルを手動で確実に傾き補正することができます。また、テキストをページ毎に画像(JPEG 等)を見ながら手動で修正することができます。

インストールは簡単で配布物をダウンロードして解凍・配置するだけですぐに利用可能となります。

Windows 環境のもと Java 利用環境があれば他の環境整備は必要としません。

PdfSlantTextCorrector は全て Java で開発されています。

定義

本ドキュメントでの語句は以下の通り定義します。

本ツール : PdfSlantTextCorrector 単体(ライブラリは含めない)を意味します。

ライブラリ : swt、pdfbox、pdfbox-tools、fontbox、commons-logging のソフトを意味します。

本ソフト : 本ツール及びライブラリを含む配布物全体を意味します。

JavaXX : OracleJDKXX、OracleOpenJDKXX、OpenJDKXX(バイナリビルドが複数ベンダよりリリースされています。)を指します。

特徴

- ・ PDF や JPEG ファイルの傾きを任意角度で確実に補正できます。
- ・ PDF では文書形式・画像形式の混合コンテンツタイプにも対応可能です。
- ・ 傾き補正操作として枠線表示、画像の拡大縮小・移動機能や補正後確認機能を実装しています。
- ・ パスワードが設定された PDF でも対応可能です。
- ・ AcrobatReader(DC)やPDF 読込ツールを起動できる機能を実装しており事前に各ページの傾きを確認できます。
- ・ 補正のため設定した角度調整データはテキスト形式で履歴として追加出力されます。そしてそのデータは初期設定値として読み込む事が可能です。
- ・ テキスト修正機能により OCR 処理で生じた誤字等を修正できるようになり、PDF テキスト検索が適正に行えるようになります。また正しいテキストも抽出できるようになります。
- ・ PDF のテキスト修正が JPEG、AcrobatReader(DC)やPDF 読込ツールの画像を見ながら修正することが可能です。
- ・ テキスト修正処理では検索機能や全置換機能も利用することもできます。
- ・ テキストボックスやTXT 形式ファイルを使用して指定ページテキストの一括修正が可能です。
- ・ 本ツールの基本的な一部属性についてはプロパティファイルを通して変更が可能です。
- ・ プロパティファイルの一部属性についてはオプション設定画面での再設定が可能です。
- ・ PDF 指定ページの読込後及び回転後画像を JPEG 形式で出力することが可能です。
- ・ PDF 指定ページの読込後及び修正後テキストを TXT 形式で出力することが可能です。
- ・ PDF 全ページあるいは指定ページの画像を JPEG 形式で出力することが可能です。

バージョン 1.5.1 改訂内容

- ・ PC 解像度に対する画面構成をより適正にするためプロパティ項目を追加しました。
- ・ PC 解像度に対応するプロパティファイルを 3 種類(HD:1280*720、WXGA-WIDE:1366*768、FULLHD:1920*1080)配布物に含めました。横サイズを基準にして利用 PC に近いファイルを採用し必要に応じて調整して下さい。

留意点

- 補正は手動で行うことが前提です。自動で補正する機能は保有していません。
- 傾き補正は対象ページを回転させることにより実現していますので、オプション選択によっては保存PDF対象ページのサイズが変わる可能性があります。（デフォルトでは元ページサイズに調整されますが、プロパティファイルのオプション選択により特定ページサイズに統一する事も可能です。）
- 回転により表示画像の鮮明度が多少低下する場合がありますが、表示画像と保存PDFの鮮明度は関係がなく表示画像が保存PDFの鮮明度に与える影響はありません。ただ、回転自体により保存PDF内補正対象ページの文字等鮮明度に若干影響を与える場合があります。
- PDFの各対象ページデータを一旦画像データに変換するためページ表示処理では時間(数秒)が多少かかります。
- JPEGファイルは1ファイル単位で操作します。
- テキスト修正はPDF画像内容に一致させることが原則です。自由に編集することはトラブルのもとであり、認めていません。またページ内の大半の文字を修正することは想定していません。あくまでOCR処理に伴う一部誤字や文字化けの修正を行うことを想定しています。
- テキスト修正できる対象ページはWindowsMicrosoft(90ms系)エンコードだけを使用するページで、Identity系エンコード(CMap埋め込み使用)等が含まれるページは対象外となります。基本的にWordやExcelから作成されたページは対象外となり(誤字等は基本的にありませんので)、スキャン(OCR処理)したPDFが対象となります。
- 表示されたページテキスト内容の多くが文字化けの場合はそのページで設定されているエンコードは対象外と判断でき、テキスト修正は不適切となります。
- プロパティファイルで、対象とするエンコード(90ms以外)を追加指定することは可能ですが、フォントメーカ等の使用許諾を本ツール利用者がとっていないエンコードは指定することはできません。

利用規約(Terms of Use)

本ツールはフリーソフト(無料ソフトという意味)です。

本ソフトの利用にあっては以下項目についての了解・遵守が必要です。

本利用規約の了解・遵守を前提として自由に使用することができます。

- 本ソフトの使用に対して、その使用目的に対する適合性について開発者(作者)及びライブラリ関連組織はいかなる保証も致しません。また使用したことによって生じた全ての障害・損害・不具合・訴訟等に関して直接・間接を問わず開発者(作者)及びライブラリ関連組織は一切その責任を負いません。全て利用者の責任においてご使用ください。
- 本ツールは全て独自開発したもので、その著作権は本ツールの開発者(作者)にあります。
- 本ツールのソースは非公開です。本ツールに対するリバースエンジニアリングは認めておりません。
- 配布物に対する改変や配布物の他ソフトへの組み込みは認めておりません。
- 配布物を本ソフト以外の用途で使用する事や他ソフトから使用することは認めておりません。
- 配布物の再配布(有料・無料を問わず)は認めておりません。

利用想定環境

本ソフトを利用するための想定環境を列挙します。

- OS : Windows
- CPU : 2GHz 相当以上

- メモリ：4GB 相当以上（本ソフトで 2GB 程度使用余裕があること。）
- 画面解像度：1280*1024 相当以上がベスト。
縦方向解像度については 768 程度でも支障なく操作可能。
- 文字の大きさ：Windows 規定設定相当以下とします。
大きな文字サイズで環境設定している場合にはコンポーネントラベル名等の一部が見えないことがあります
が、プロパティファイルで指定することによりコンポーネントの文字サイズを変更することは可能です。

Java 環境の整備

本ソフトでは Java を使用するため、利用者 PC 環境に Java がインストールされていることが前提となります。

Java は Java10 までは無償で商用利用も可能でしたが、Java11 (OracleJDK)以降は有償(個人での利用や開発での利用などは無償)となりましたが、現時点(2021. 11)での最新版 Java17 (OracleJDK LTS 版)では再度無償となりました。現在 Java は基本半年間隔でリリースされサポート期間も半年となっており(3 年毎には長期サポート版はリリースされます)、またライセンス関係や有償/無償切替問題もあり Java (JDK や JRE)は配布物に含めていません。

本ソフトは Java8、Java10、Java17 で稼働できることを確認しており、Java17 は OracleJDK17 や OpenJDK17 (無償・商用利用可)として複数ベンダからバイナリビルドを入手して利用することが可能となっています。

今回テストで使用した EclipseTemurin(旧 AdoptiumOpenJDK : OpenJDK)17 のインストール手順を参考までに以下に記述します。

- ・以下の URL からインストーラ (msi ファイル：約 164MB) をダウンロードします。
<https://adoptium.net/>
- ・msi をダブルクリックして EclipseTemurin17 (JDK) をインストールします。
(カスタムセットアップ画面で各項目を設定しなければ(×にする)パス等の PC 環境が変更されることはありません。)

このように Java17 は簡単にインストールすることができ、Java8、Java11 についても EclipseTemurin からダウンロード可能となっています。

なお、本ソフト利用者自身が、自身の PC 環境で、第三者を関係させず Java (JDK) を利用する分には問題ありませんが、その他の利用・活用する場合には該当ライセンスや利用規約をよく読んで対応する必要があります。

ライブラリの配備と使用

本ソフトでは swt、pdfbox、pdfbox-tools、fontbox、commons-logging の 5 つのソフト(ライブラリ：jar 形式)を使用します。このライブラリは jar ディレクトリ配下の lib (Java8, Java10 用) と lib17 (Java17 用) ディレクトリに格納されています。

Java17 を使用する時は lib ディレクトリのセットを lib17 ディレクトリのセットで置き換えるか、起動用パッチファイル(*)の 2 ヶ所(lib→lib17)を手動で変更することで対応可能です。

この起動用パッチファイルの変更は配布物(batchupdater ディレクトリ)に含まれている batchupdater.bat を実行することで可能です。batchupdater.bat では Java バージョンの選択とライブラリの選択が可能です。

batchupdater の利用方法は以下の通りです。

- ・batchupdater.bat を起動すると「バッチ内 Java パス変更画面」が表示されます。
- ・「Java 検索」ボタンを押下すると PC 内で利用可能な Java がラジオボタン形式で表示されます。
- ・使用する Java バージョンを選択します。

- ・次に使用するライブラリの種類を選択します (Java バージョンに合わせて下さい)。
- ・「バッチ更新」ボタンを押下すると起動用バッチファイルが更新されます。
- ・「完了」ボタンを押下して終了します。

留意事項は以下の通りです。

- ・batchupdater 起動は「JAVA_HOME」環境変数が設定されているという前提になっています。「JAVA_HOME」環境変数が設定されていない場合は batchupdater.bat の jbindir 変数をコメント化設定行を参考に Java パスを適切に設定して下さい。
 - ・PC 内 Java バージョンの検索は C:\¥pdfstcorrector¥batchupdater¥javadata.txt をもとに行われ、javaPath キーで指定したディレクトリでの Java バージョンと、「JAVA_HOME」環境変数での Java バージョンが検索されます。
 - ・Java バージョン検索用ディレクトリや変更起動用バッチファイルは画面から選択 (1 つ) することも可能です。
 - ・必要に応じて javadata.txt データの値を変更することは可能ですが、データ順序の変更はできません。
- (*) : C:\¥pdfstcorrector¥batch ディレクトリにある pdfstcorrector.bat 及び pdfstcorrector_vbs.bat。

インストール

本ソフトのインストール作業について説明します。

なお、Java 利用環境については既に整備されているものとします。

またプロパティファイルについては PC 解像度に合わせて選択し必要に応じて調整して下さい。

- ・インストールは基本的にダウンロードした zip ファイルを解凍・配置するだけです。
- ・C ディレクトリ配下に pdfstcorrector ディレクトリがあればこのディレクトリ名を変えて一旦構成・内容を保持しておいて下さい。
- ・ダウンロードした zip ファイルを解凍して C ディレクトリ直下に配置して下さい。
- ・「C:\¥pdfstcorrector¥batch」内の起動用バッチファイル (pdfstcorrector.bat 及び pdfstcorrector_vbs.bat) 内に設定されている jbindir 変数を変更して利用者環境の Java パスを設定して下さい。パスに半角ブランクを含む場合は” ” で囲んで下さい。(例 : set jbindir="C:\¥Program Files¥Java¥jre-10.0.2¥bin") (*1)
- ・Java17 を使用する場合はライブラリを置換えるか、起動用バッチファイル内のライブラリパスを変更して下さい。(*1) (*2)
- ・AcrobatReader (DC) あるいは PDF を読込できるツールがある場合、そのパスをプロパティファイルで設定して下さい。(キー : acrobat_reader_path など)
- ・起動は「C:\¥pdfstcorrector¥batch」ディレクトリに配置されている pdfstcorrector.bat あるいは pdfstcorrector.vbs のどちらかをダブルクリックして行います。

(vbs 起動の場合 DOSWindow が表示されません。)

(*1) : batchupdater.bat 実行で対応することも可能です。

(*2) : 「ライブラリの配備」参照。

アンインストール

- ・「C:\¥pdfstcorrector」ディレクトリを含め配下を全て削除して下さい。
- ・レジストリは使用しておりません。

開発及びテスト環境

本ソフトの開発及びテスト環境は以下の通りです。

PC : CPU:1.8GHz, メモリ:4GB, ディスプレイ:1366×768 14 インチ WXGA ワイド

OS : Windows 8.1 64 ビット

Java : Java8、Java10、Java17

SWT ライブラリ : SWT4.9、SWT4.21 Windows 64bit 版

PDFBox ライブラリ : PDFBox2.0.14、PDFBox2.0.24 版

CommonsLogging ライブラリ : CommonsLogging1.2 版

配布物

配布物には以下のツールやライブラリ (jar ファイル形式) が含まれています。

ライブラリはいずれも改変せずそのまま使用しています。

pdfstcorrector : 本ツールの Java プログラム (クラスファイル) が格納されています。

swt : Java 用 ウィジェット・ツールが格納されており、開発・管理元は Eclipse Foundation で、Eclipse Public License でライセンスされています。

pdfbox(*) : PDF を操作するためのツールが格納されており、開発・管理元は Apache Software Foundation で、Apache License Version 2.0 でライセンスされています。

commons-logging : 主としてロギングするためのコンポーネントが格納されています。開発・管理元は Apache Software Foundation で、Apache License Version 2.0 でライセンスされています。

(*)pdfbox-tools、fontbox を含めます。

ライブラリのライセンス関係ファイルは配布物内の license ディレクトリに格納されています。

基本操作手順

傾き補正とテキスト修正を連続処理してまとめて PDF 保存することはできません。

両処理を行う場合は、まずは傾き補正をして一旦 PDF を作成し、それを使用してテキスト修正を行う手順となります。傾き補正をしない場合は「読込」だけを行いテキスト修正に進むことが可能です。

○基本的な傾き補正の操作手順は以下の通りです。3、4、6、7 は PDF ファイル複数ページ処理の場合となります。コンポーネントの左側→右側に移動しながら操作していきます。

1. 対象ファイルの形式を「PDF/JPEG 選択」で選択する。
2. 「読込」で対象ファイルを選択し読込みます。(パスワード付の場合パスワード設定も行います。)
3. (PDF 複) ページ番号指定機能を使用して対象ページを指定します。
4. (PDF 複) 「表示」で指定ページの画像変換及び表示を行います。(多少時間がかかります。)
5. 回転、拡大、移動機能を使用して傾き補正のための回転角度を決定します。
6. (PDF 複) 「更新」でページ番号と回転角度を更新保存します。
PDF 更新時画像 (JPEG) 出力オプションが 2 の場合ページの回転後 JPEG 画像が出力されます。
7. (PDF 複) 3. ~6. を必要に応じて繰り返します。
8. 「保存」で設定・更新したデータにもとづいて回転処理した PDF or JPEG ファイルを保存します。
9. 他ファイルを実行する場合は 2. から繰り返します。
10. 「完了」でツールの処理を終了させます。

なお、キャンバスに描写される画像と保存される PDF との鮮明度には関係性がなく、表示されている

画像がやや不鮮明でも適切な傾き補正角度を決定できれば保存PDFには影響ありません。

○基本的なテキスト修正の操作手順は以下の通りです。1. ～2. は傾き補正操作手順と同じです。

1. 対象ファイルの形式を「PDF/JPEG 選択」で選択する。(PDF を選択)
2. 「読込」で対象ファイルを選択し読込みます。
3. 「T 更新」でテキスト更新起動画面を起動します。
4. テキスト更新起動画面で対象 PDF ファイルと対象ページを指定し、修正する場合は「テキスト更新」を押下します。右側の「テキスト表示」では下部テキストボックスにテキスト内容を表示することができ事前確認に利用することができます。(オプション設定により「テキスト表示」押下時に修正前テキスト内容を出力することが可能。)
5. 起動したテキスト更新画面でテキストを修正します。PDF 画像を表示する場合は「PR」あるいは「表示」を押下します。「PR」は AcrobatReader 等が起動し、「表示」は JPEG 画像が表示されます。「PR」の場合は画面右側余白にサイズ・位置を調整して表示します。JPEG 画像は上下左右に移動、拡大縮小させることができます。
6. 表示されたテキストはトークン単位で改行されています。テキスト修正前であれば「改行挿入」や「改行削除」により改行位置を変更することができます。なお、改行設定は表示上（見栄え）のことであり、PDF テキスト内容には影響しません。設定した改行位置は保存することができます。改行は改行文字列（デフォルト値（5 文字）：%¥n%+改行文字）で表示されています。なお、この改行処理の必要性は特段あるわけではなく、ユーザの見やすさのための機能です。
7. テキストボックスや TXT 形式ファイルから一括修正する場合はテキスト更新画面の「TBOX 更新」、「TBOX 再更新」（2 回目以降の更新時に使用）や「ファイル更新」ボタンを操作して行います。ただし、挿入、更新、削除処置の前に行う必要があります。また、挿入、更新、削除操作を行うと「TBOX 更新」、「TBOX 再更新」や「ファイル更新」ボタンは非活性となります。
8. その後必要であれば挿入、更新、削除処理を行い、該当ページの全修正が終了した後、「更新確定」を押下し、画面を終了する場合は「完了」を押下します。
挿入、更新、削除処理は基本的にマウスカーソルでキャレット or レンジで位置や文字列を選択し、「選択」ボタンを押下して確定し、更新データテキストで変更文字列を設定し、「右挿入」、「左挿入」、「更新」、「削除」ボタンを押下して行います。
(「更新確定」は複数回操作が可能で、修正は累積されていきます。)
テキストファイルには改行文字列が適正数含まれている必要があります。
テキスト更新完了時テキスト出力オプション指定によって出力されたテキストファイルをこの「ファイル更新」処理の元ファイルとして利用することができます。
9. 他ページのテキスト修正がある場合は一旦テキスト更新画面を終了し、4. ～8. を繰り返します。
テキスト更新時テキスト出力オプションが 2 の場合「完了」押下後修正後ページ内容を TXT 形式で出力します。
10. PDF ファイルにおける全てのページテキスト修正が完了した場合はテキスト更新起動画面の「PDF 保存」を押下します。
11. 他 PDF ファイルのテキスト修正がある場合は 2. ～10. を繰り返します。
他 PDF ファイルがパスワード付 PDF で、傾き補正「読込」で設定したパスワードと異なる場合はエラーとなりますので、一旦テキスト更新画面をクローズし、PDF 傾きテキスト修正基本画面の「読込」操作で再設定して下さい。

12. 全 PDF ファイルのテキスト修正が完了した場合にはテキスト更新起動画面の「完了」を押下します。

画面説明

本ツールは主に傾き補正用 1 画面とテキスト修正用 2 画面の計 3 画面と、サブ画面であるオプション設定画面と TEXTBOX 更新画面、JPEG 変換ページ設定画面を含めた合計 6 画面で構成されています。

画面内コンポーネントの説明を以下に記述します。

操作は基本的に左→右コンポーネント、上→下コンポーネントの順にしたがって操作していきます。

基本画面(PDF 傾きテキスト修正基本画面)：

コンポーネント名	タイプ	説明
PDF/JPEG 選択	ラジオ	補正するファイルのタイプを選択します。
「枠線」	チェック	枠線を表示するかのオプションチェックボタン。 チェックすることにより枠線が表示されます。 枠線数はプロパティファイルで変更可能です。
「OPT」	ボタン	プロパティファイルでの設定値の一部（8 値）を画面から指定できます。 プロパティファイル設定値より優先されます。 読込表示オプション、回転後の画像表示・保存オプション、回転後の PDF 保存オプション、90 度倍数の回転角度での保存オプション、PDF 更新時画像 (JPEG) 出力オプション、テキスト更新完了時テキスト出力オプション、PDF 表示時画像 (JPEG) 出力オプション、テキスト表示・更新操作時テキスト出力オプション。 前 4 値を変更した場合、「読込」操作からの再操作となります。 「プロパティ」ボタンは各値をプロパティファイル設定値に設定します。
「読込」	ボタン	対象ファイルを読込みます。ファイル選択ダイアログがオープンします。 格納ページ数が 1 の場合そのまま画像データに変換して画面表示します。 複数ページの場合一旦合計ページ数が表示され、その後「表示」機能を使用して選択ページを画面表示します。 読込が終了すると読込んだページを引継いで回転処理するかのダイアログが表示されます。「はい」押下の場合は引継いで回転処理操作 (or テキスト更新) を行い、「いいえ」押下の場合は回転処理操作 (or テキスト更新) の前に読込んだページの JPEG 変換を行うため JPEG 変換ページ設定画面が開きます。 ***パスワードについて*** パスワードが設定されている場合にはパスワード入力用ダイアログがオープンします。ユーザ/権限パスワードを PDF ファイルの設定状態に合わせて（原則現設定値に一致させて）入力します。両方設定している場合は両方入力する必要があります。 読込処理は基本的にユーザパスワードを使用して行われます。 保存処理は入力設定したパスワード（1 個 or 2 個）で行われ、保存された PDF は入力設定したパスワードでオープン/編集する必要があります。
「PR」(*)	ボタン	プロパティファイルでの設定にもとづいて PdfReader (AcrobatReader 等) を起動

		し、選択ファイルを表示します。起動後は本ツールとは非同期となります。 傾きのあるページを検索する場合などで使用します。
「←」(*)	ボタン	ページ番号を 1 減少させます。
ページ番号指定(*)	テキスト	表示ページ番号を指定します。「←」「→」で増減したページ番号はこのテキストで表示されます。読込後は 1 に設定されています。 enter キーで表示処理を実行します。
「→」(*)	ボタン	ページ番号を 1 増加させます。
更新モード	ラベル	PDF の更新・保存モードを表示しています。 PDF の回転角度を更新する方法には 2 種類あります。 T モード：表示処理せずに更新する方法で、「ページ番号指定」テキストと「回転角度指定」テキストの値を使用して更新処理します。 ただし読込後「表示」処理すると D モードに切り替わります。 D モード：表示処理しながら更新する方法で、「表示ページ番号」ラベルと「合計回転角度」ラベルの値を使用して更新処理します。 T モードは PDF→画像データの変換時間が省略できるメリットがありますが、予め補正すべきページ番号と補正角度がわかっている必要があります。 基本的には D モードで操作処理していきます。
「表示」(*)	ボタン	対象ファイルが PDF で複数ページの場合、ページ番号指定で設定されたページを画面表示します。表示ページ番号ラベルにページ番号が表示されます。 PDF 表示時画像 (JPEG) 出力オプションが 2 の場合、該当ページの回転前 JPEG 画像が出力されます。
表示ページ番号	ラベル	画面表示されているページ番号を表示します。
「初期」	ボタン	回転処理、拡大処理、移動処理をクリアして読込直後状態にします。 ページ番号:1、回転角度:0、拡大率:1.0、移動:(0,0)となります。
合計回転角度	ラベル	合計回転角度を表示します。 次項目の回転角度指定では現表示画像を何度回転させるかを指定します。
回転角度指定	テキスト	現表示画像への回転角度 (単位: 度) を指定します。 右回転が正值で左回転が負値となります。次項目「回転」を連続して押下した場合指定した角度で連続追加回転していきます。(*1) 参照。 enter キーで回転処理を実行します。
「回転」	ボタン	現表示画像を指定した角度で回転します。 表示画像やサイズが変更されるため一旦スライダー位置は 0,0 に戻されます。 回転済画像を元のスライダー位置に移動する場合は継続操作として「移動」を押下すれば移動させることができます。
拡大率指定	テキスト	現表示画像に対する拡大率を指定します。 >1.0 で拡大、<1.0 で縮小となります。(*1) 参照。

		enter キーで拡大処理を実行します。
「拡大」	ボタン	<p>現表示画像を指定した拡大率で拡大・縮小します。</p> <p>表示画像やサイズが変更されるため一旦スライダー位置は 0,0 に戻されます。</p> <p>拡大済画像を元のスライダー位置に移動する場合は継続操作として「移動」を押下すれば移動させることができます。</p>
「移動」	ボタン	<p>現表示画像をスライダーバー位置まで縦・横移動します。</p> <p>本ツールでの画像移動はスライダーとマウスホイールで行えます。</p> <p>ただし、スライダー・マウスホイール操作での処理発生タイミングは指定間隔でしか発生しないので、この「移動」ボタンにより任意のタイミングで処理を強制的に実行できるようにしています。</p> <p>イベント間隔はプロパティファイルで指定することができます。</p> <p>また、「回転」や「拡大」操作では一旦スライダー位置は 0,0 になるため、操作直後この「移動」ボタンを押下すれば元スライダー位置に移動します。</p>
「更新」	ボタン	<p>対象ファイルが PDF ファイルの場合に使用します。</p> <ul style="list-style-type: none"> PDF 内指定ページを回転処理し、指定ページのページ番号と合計回転角度を設定保存します。設定保存した補正用データ（ページ番号、合計回転角度）は「確認」操作で確認することができます。 PDF 更新時画像 (JPEG) 出力オプションが 2 の場合、該当ページの回転後 JPEG 画像が出力されます。 <p>—以下の操作・処理は PDF が複数ページ構成の場合において有効です。—</p> <ul style="list-style-type: none"> 読込操作直後に更新押下操作を行うことにより補正用データ（補正用角度調整データ）を初期値として読込むことができます。読込形式は保存処理で履歴として出力される補正用データと同じで DateTime～Dataxxx までとなり、履歴出力ディレクトリに history_input.txt ファイルとして配置します。 「表示ページ番号」ラベルと「ページ番号指定」テキストの値が異なる場合は警告ダイアログが表示され、どちらのページ番号を使用するかを選択します。「ページ番号指定」テキストを選択した場合、保存される回転角度は「回転角度指定」テキスト値となります。（この処理は確認処理で補正後画像を表示確認できない場合があるので推奨はしません。）
確認ページ番号(*)	テキスト	<p>確認するページ番号を入力します。</p> <p>0：連続確認処理を開始します。「確認」を連続押下して更新保存されたページ番号と合計回転角度を順次確認します。最下部のメッセージテキストにデータが表示され、更新保存角度分回転された画像が画面表示されます。上述の「ページ番号指定」テキスト優先処理をした場合画像が表示されない場合があります。</p> <p>x：連続確認処理を強制終了します。</p> <p>その他任意数値：指定したページの保存データ及び画像を確認できます。</p>
「確認」(*)	ボタン	更新保存されたデータを確認するために使用します。

「保存」	ボタン	更新保存されたデータ群をもとに PDF (単ページ) / JPEG を回転させて保存します。元ファイル名最後尾に追加文字列を付加して補正後ファイルを保存します。追加文字列はプロパティファイルで変更できます。設定された補正用データ (補正用角度調整データ) は履歴として追加出力されます。
「T 更新」	ボタン	テキスト修正を行うためのテキスト更新起動画面を起動します。
「完了」	ボタン	本ツール処理を終了します。 ウィンドウ右上の X で完了させることはできません。
(*)		PDF が 1 ページ構成の場合は非活性状態となります。
(*1)		<p>値の入力には 2 種類の方法があります。</p> <p>数値だけの場合：現表示画像に対する処理値となります。</p> <p>特殊文字 (2 文字) + 数値：指定操作後の最終画像に対する処理値となります。</p> <p>特殊文字は以下の 4 種類指定できます。</p> <p>rd/rt/sc/mv：読込/回転/拡大/移動</p> <p>例 1：回転指定 sc5→拡大処理後の最終画像を 5 度右回転</p> <p>例 2：拡大率指定 rd1.1→読込処理後の最終画像を 1.1 倍に拡大</p> <p>後者の入力 (特殊文字使用) 方法は複数回使用すると混乱する可能性がありますので複数回使用については推奨しません。</p>

テキスト更新起動画面：

コンポーネント名	タイプ	説明
「パス選択」	ボタン	対象とするパス (ディレクトリ) を選択します。ファイル選択ダイアログがオープンしますので対象とするファイルを選択して下さい。パスが決定され、そのパスに含まれている PDF ファイルをファイル名一覧コンボに設定します。 画面起動時には傾き補正画面で読み込んだパスが既に設定されています。
ファイル名一覧	コンボ	テキスト修正する PDF ファイルを選択します。 画面起動時には傾き補正画面で読み込んだファイルが選択されています。
「ファイル選択」	ボタン	選択した PDF ファイルを決定し、ページ番号一覧コンボにページ番号を設定します。
ページ番号一覧	コンボ	設定されているページ番号からテキスト修正するページを選択します。
「テキスト更新」	ボタン	対象 PDF ファイル及びページを決定し、該当ページを解析し、テキスト更新画面を起動し、ページのテキスト内容やボタン類等を表示します。 テキスト表示・更新時テキスト出力オプションが 2 の場合、修正前ページテキスト内容が TXT ファイル形式で改行文字列込みで出力されます。
「テキスト表示」	ボタン	対象 PDF ファイル及びページを決定し、該当ページを解析し、下部テキストボツ

		<p>クスに修正前テキスト内容を表示します。</p> <p>テキスト表示・更新時テキスト出力オプションが2の場合、修正前ページテキスト内容がTXT ファイル形式で改行文字列込みで出力されます。</p>
「更新」	ボタン	設定されているパスに含まれる PDF ファイルを再検索し、ファイル名一覧コンボに設定します。
「PDF 保存」	ボタン	操作された更新(修正)データにもとずいて、更新した新規の PDF ファイルを作成します。新規 PDF ファイル名は元 PDF ファイル名に指定文字列(デフォルト : _textupdate)を後尾に追加したものととなります。既存ファイルがある場合は更に後尾に日時を追加することができます。
空対応無効	チェック	新規 PDF ファイルを作成する際トークン (行) が空の場合は全角ブランク 1 文字を自動的に設定するようになっています (後日該当部分に文字を挿入する事を想定して)。チェックした場合この機能を除外することになります。ただしチェックすることは以後利用の観点から推奨しておりません。
「完了」	ボタン	テキスト更新起動画面を終了します。

テキスト更新画面：

コンポーネント名	タイプ	説明
テキスト内容表示	テキスト	選択したページのテキストを表示します。
トークン改行無効	チェック	基本的にテキストはトークン単位で改行が追加されて表示されますが、チェックをして再読込した場合には改行されずに表示され、ページのコンテンツ自体の内容を確認できます。確認後はチェックを外して再読込をする必要があります。
「再読込」	ボタン	選択しているページのテキストを再読込をします。ただし、それまでに操作した更新情報や保存情報がクリアされるので注意が必要です。
「再表示」	ボタン	誤って「テキスト内容表示」テキストを直接編集し、表示内容が変わってしまった場合、表示内容を最終更新後の状態に戻します。
「改行挿入」	ボタン	表示テキストのカーソル位置に改行文字列を挿入します。 改行挿入は右挿入/左挿入/更新/削除処理の前に実施しておく必要があります。 (実際の PDF テキストに挿入するわけではありません。)
「改行削除」	ボタン	表示テキストのカーソル位置の改行文字列を削除します。 改行削除は右挿入/左挿入/更新/削除処理の前に実施しておく必要があります。 (実際の PDF テキストを削除するわけではありません。)
「保存」	ボタン	現在の改行位置を保存します。改行挿入、改行削除しなければ特に操作の必要はありません。
「保存読込」	ボタン	保存した改行位置情報を読込ます。ただし、それまでに操作した更新情報や保存情報がクリアされるので注意が必要です。

「選択」	ボタン	カーソルで指定した位置（キャレット）や範囲（レンジ）を更新処理のため決定します。
「複写」	ボタン	指定したレンジの文字列を更新データテキストにコピーします。
「行削除」	ボタン	指定したレンジの行（複数行可能）を削除します。
「選択中止」	ボタン	選択したキャレット位置やレンジをクリア（無効に）します。
「TBOX 更新」	ボタン	<p>ページ全体の一括更新（修正）をサブ画面のテキストボックス設定内容を使用していきます。適切な改行文字列数が含まれている必要があります。</p> <p>挿入、更新、削除操作より先に操作する必要があります。</p> <p>テキストボックスを使用した更新は「TBOX 更新」及び「TBOX 再更新」ボタンを使用して複数回繰り返して行うことが可能です。</p> <p>一度「TBOX 更新」操作を行うと「TBOX 更新」は不活性となり、「TBOX 再更新」が活性化されます。</p>
「TBOX 再更新」	ボタン	「TBOX 更新」で更新された内容をさらに更新する時に使用します。
「ファイル更新」	ボタン	<p>ページ全体の一括更新（修正）をテキストファイルより行います。</p> <p>適切な改行文字列数が含まれている必要があります。</p> <p>挿入、更新、削除操作より先に操作する必要があります。</p>
更新データ	テキスト	更新（修正）するデータ（文字列）を設定します。
「左挿入」	ボタン	<p>指定したキャレット位置に更新データを挿入します。</p> <p>改行文字列の直前直後の場合には左側トークン（文字列）に挿入します。</p>
「右挿入」	ボタン	<p>指定したキャレット位置に更新データを挿入します。</p> <p>改行文字列の直前直後の場合には右側トークン（文字列）に挿入します。</p>
「更新」	ボタン	指定したレンジの文字列を更新データに変更します。
「削除」	ボタン	指定したレンジの文字列を削除します。
「←」	ボタン	現検索結果より 1 つ前のレンジ(文字列)へカーソルを移動します。
「前検索」	ボタン	1 つ前の検索ワードで再検索します。
「検索」	ボタン	カーソルのレンジ(文字列)あるいは更新データテキストの文字列を使用して検索します。
「→」	ボタン	現検索結果より 1 つ後のレンジ(文字列)へカーソルを移動します。
「全置換」	ボタン	「選択」で指定したレンジの文字列を更新データテキストの文字列に全て一括置き換えます。
「PR」	ボタン	現在のページ画像を PdfReader（AcrobatReader 等のパスをプロパティファイルで指定）を起動させて表示させます。
「表示」	ボタン	現在のページを JPEG 画像で画面右側空白領域に表示させます。

「+」 「-」	ボタン	表示した JPEG 画像を拡大・縮小します。
「top」	ボタン	表示した JPEG 画像の最上部を表示します。
「↑」 「↓」	ボタン	表示した JPEG 画像を上下へ移動します。
「←」 「→」	ボタン	表示した JPEG 画像を左右へ移動します。
「 ←」 「→ 」	ボタン	表示した JPEG 画像の左端部あるいは右端部を表示します。
「更新確定」	ボタン	更新（修正）した操作（データ）内容を確定します。
「完了」	ボタン	テキスト更新画面を終了します。 テキスト更新完了時テキスト出力オプションが2の場合、修正後ページテキスト内容がTXT ファイル形式で改行文字列込みで出力されます。

オプション設定画面：

コンポーネント名	タイプ	説明
「プロパティ」	ボタン	各オプション値をプロパティファイル値に設定します。
「設定」	ボタン	各オプション値を画面で選択したオプション値に設定します。 オプション設定画面は終了します。
「完了」	ボタン	オプション設定画面を終了します。

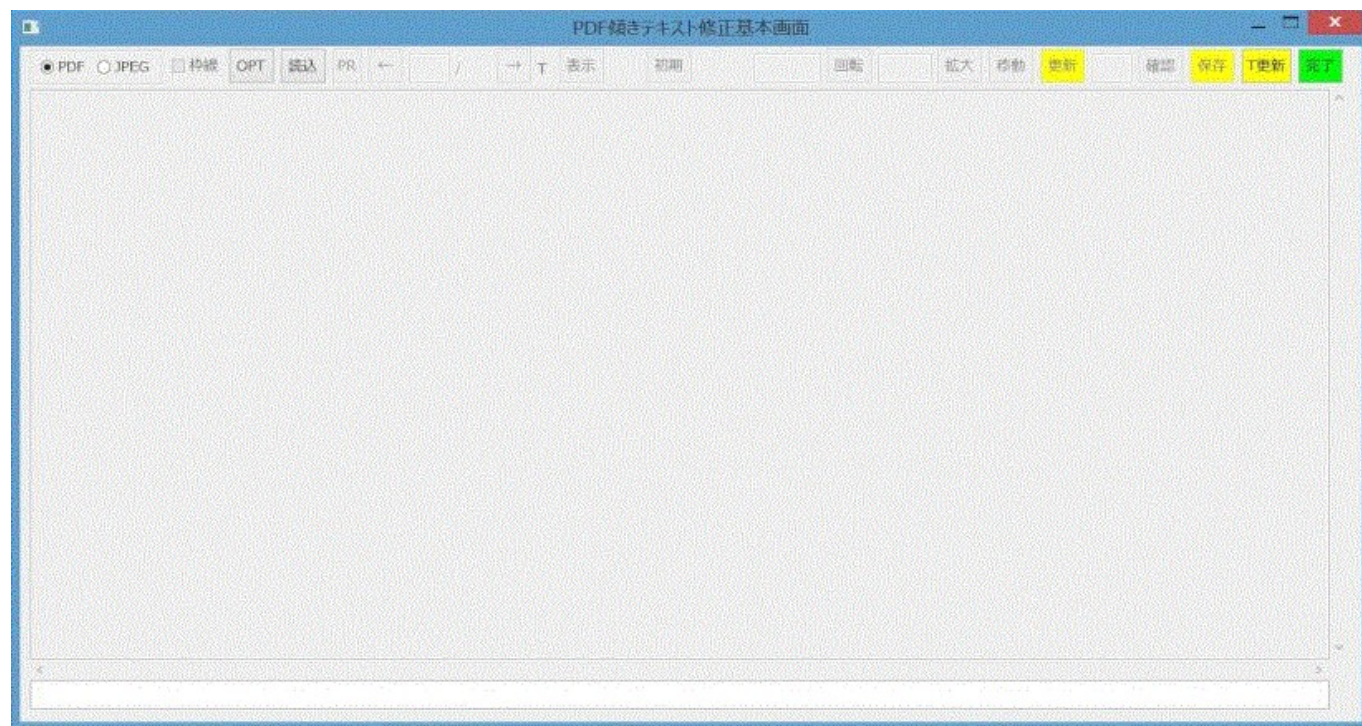
TEXTBOX 更新画面(TEXTBOX 再更新画面も同様)：

コンポーネント名	タイプ	説明
「確定」	ボタン	テキストボックスで設定した内容に対象ページのテキストを更新(再更新)設定します。TEXTBOX 更新画面は終了します。
「キャンセル」	ボタン	テキストを更新せずに TEXTBOX 更新画面を終了します。

JPEG 変換ページ設定画面：

コンポーネント名	タイプ	説明
変換オプション	ラジオ	全ページを変換対象とするか部分ページを対象とするかを選択します。
ページ番号設定	テキスト	部分ページを対象とする場合このテキストで対象とするページを設定します。
「変換」	ボタン	設定されたページの JPEG 変換処理を開始します。 一定時間(10 秒)毎に処理経過をダイアログで表示し処理継続の確認を行います。 JPEG 変換処理は完了しないので、継続して JPEG 変換操作が可能です。
「完了」	ボタン	JPEG 変換処理を完了します。

基本画面



プロパティファイル説明

プロパティファイルは「キー=値」の形式で記述された Java のテキストファイルで、規定値や変更の可能性がある設定値などを格納します。

本ツールで使用しているプロパティファイルは以下の 1 種類です。（prop ディレクトリ内に配置。）

- ・pdfstcorrector.properties：本ツール実行のために必要とする基本的な規定値・設定値格納ファイル。

プロパティファイルにはユーザが変更可能な項目と変更できない項目が識別できるようにコメント追加しています。更新欄では「-」がほとんど更新の必要のないキー、「△」が状況により更新しても良いキー、「○」が更新度の高いキーを意味しています。

1.5.1 版で追加したキーは以下の通りです。

optionshell_position、optionshell_size、jpegshell_position、jpegshell_size、
textupdate_text_comp、option_margin、jpeg_margin

PC 解像度によって変更しているキーは以下の通りです。

font_height、shell_size、button_comp、radiobutton_comp、optionshell_size、jpegshell_size、
pageshell_size、textupdate_button_comp、updateshell_size、textboxshell_size、
textupdate_canvas_margin

更新	キー	説明	標準値
		<*****ユーザが変更可能データ*****>	
○	acrobat_reader_path	PdfReader (AcrobatReader 等) 起動用ファイルパス。 (例) C:\Program Files (x86)\Adobe\Acrobat Reader DC\Reader\AcroRd32.exe	-
○	acrobat_reader_param	PdfReader (AcrobatReader 等) 起動用コマンドパラメータ。 (例) /n /o	-
△	pdf_reader_cmd	PdfReader (AcrobatReader 等) 起動用 DOS コマンド。 PDF 閲覧を特別なツールで起動したい場合に DOS コマンドを直接設定して起動します。acrobat_reader_path より優先されます。 PDF ファイルパスは%pdffilepath%で設定。自動的に選択したパスに置換されます。 (例) "C:\Program Files (x86)\Adobe\Acrobat Reader DC\Reader\AcroRd32.exe" /n /o "%pdffilepath%"	-
△	extension_name	保存用新規ファイル名として指定元ファイル名最後尾に付加する文字列 (例：abc.pdf の補正後保存ファイル名はabc_rotation.pdf)。	_rotation
△	grid	枠線の数(x;y)。	20;20
-	read_display_option	読込画像のコンテンツ表示オプション。 1:キャンバスに画像全体が入るようにリサイズ(2つ目の値は1.0)。 2:画像短辺がキャンバスに入るようにリサイズ(長辺はキャンバスサイズ外となる可能性有)。 2つ目の値(float)：キャンバスの何割に表示するか縦横ファクター。	2;0.8
-	pdf_save_option	回転後の PDF 保存オプション。	2;-1

		<p>1:縮小してページ全体を表示できるように保存(表示内容が小さくなる可能性有。2つ目の値は0)。</p> <p>2:縮小せずに保存あるいは指定ページのサイズに調整して保存(2つ目の値が0の場合ページサイズが大きくなります。)</p> <p>2つ目の値:-1=元ページサイズで保存、0=サイズ調整なし、>0=指定したページのサイズで保存</p>	
-	jpeg_save_option	<p>回転後の JPEG 表示・保存オプション。</p> <p>以下2つのオプションがあり、標準的にはオプション1を選択。</p> <p>写真画像のように表示空き領域を発生させたくない場合はオプション2を選択。ただし、回転角度を大きくすると表示領域がかなり小さくなっていくのでオプション2では注意必要。</p> <p>1:縮小して画像全体を表示できるように保存(表示空き領域が発生)。</p> <p>2つ目の値:0</p> <p>2:表示空き領域が発生しないようにサイズカット調整して保存。</p> <p>(ただし指定回転角度以上の場合自動的にオプション1に切り替わります。)</p> <p>2つ目の値:オプション1に切り替わる角度(デフォルト:10度)</p>	1;0
△	updatejpeg_save_option	<p>PDF 更新時画像 (JPEG) 出力オプション</p> <p>1:出力しない。</p> <p>2:更新操作時に毎回出力する。</p>	1
△	updatetext_save_option	<p>テキスト更新完了時テキスト出力オプション</p> <p>1:出力しない。</p> <p>2:更新完了時に毎回出力する。</p>	1
△	displayjpeg_save_option	<p>PDF 表示操作時での回転前画像 (JPEG) 出力オプション</p> <p>1:出力しない。</p> <p>2:表示操作時に毎回出力する。</p>	1
△	displaytext_save_option	<p>テキスト表示・更新操作時での修正前テキスト出力オプション</p> <p>1:出力しない。</p> <p>2:表示・更新操作時に毎回出力する。</p>	1
△	rotation90_option	<p>90 度倍数の回転角度での保存オプション。</p> <p>1: ページ(用紙)は回転せず保存。</p> <p>2: ページ(用紙) も一緒に回転させて保存。</p>	1
-	textupdate_extension_name	<p>テキスト更新後の新規ファイル名として指定元ファイル名の最後尾に付加する文字列。</p>	_textupdate
-	permission_encoding_name	<p>許可するエンコーディングタイプ。</p> <p>注意: CMap を使用するエンコーディングは指定不可です。</p> <p>指定エンコーディングでページ内容表示処理を行って、エラーや文字化けが発生する場合は指定不可となります。</p>	90MS
-	save_keyword	<p>前の検索キーを表現する文字列。</p>	%s%

		<*****必要性が生じた場合に限ってユーザ変更が可能なデータ*****>	
△	font_height	コンポーネント文字サイズ	9
-	font_name	コンポーネント文字フォント	メイリオ
-	shell_position	シェル(画面)の表示位置(X;Y)	30;5
△	shell_size	シェル(画面)のサイズ(X;Y)	1320;700
-	button_comp	ボタンコンポーネントの属性値(幅;高さ)	45;35
-	radiobutton_comp	ラジオボタンコンポーネントの属性値(radio;pdf;jpeg コンボ幅)	125;45;75
-	text_comp	テキストコンポーネントの属性値(幅;高さ)	70;28
-	pdf	PDF 設定値(最大ページ数;解像度 dpi デフォルト値)	450;300
-	slider_event	スライダーコンポーネントのイベント発生間隔(0-100 移動間での間隔)。間隔を標準より短くしても画像移動処理に時間がかかるため描写に遅れが生じ効果はあまり期待できない。	5
-	slider_mousewheel_increment	マウスホイールのクリックあたりのスライダー移動間隔(0-100 移動間での間隔)。間隔を標準より短くしても画像移動処理に時間がかかるため描写に遅れが生じ効果は期待できない	5
-	canvas_backcolor	キャンバスの背景色(表示 RGB3 値;保存 RGB3 値 -1 はデフォルト色指定)。	-1;-1;-1;-1;-1;-1
-	optionshell_position	オプション設定シェル(画面)の表示位置	120;20
△	optionshell_size	オプション設定シェル(画面)の表示サイズ	1000;630
-	jpegshell_position	JPEG 変換ページ設定シェル(画面)の表示位置	120;20
△	jpegshell_size	JPEG 変換ページ設定シェル(画面)の表示サイズ	510;420
-	pageshell_position	PDF ファイル・ページ選択シェル(画面)の表示位置	100;20
△	pageshell_size	PDF ファイル・ページ選択シェル(画面)の表示サイズ	650;650
-	textupdate_button_comp	テキスト更新用ボタンコンポーネントの属性値(幅;高さ)	90;35
-	textupdate_text_comp	テキスト更新用テキストコンポーネントの属性値(幅;高さ)	70;30
-	return_string	改行文字列(%で囲む)	%¥¥n%
-	updateshell_position	テキスト更新シェル(画面)の表示位置	10;-40
△	updateshell_size	テキスト更新シェル(画面)の表示サイズ	1170;700
-	textboxshell_position	テキストボックス更新シェル(画面)の表示位置	200;-40
△	textboxshell_size	テキストボックス更新シェル(画面)の表示サイズ	700;700
-	nopermission_encoding_name	許可しないエンコーディングタイプ。 指定した値のエンコーディングタイプがページに含まれる場合はテキスト更新はできません。	IDENTITY
-	jpeg_move_interval	JPEG 画像移動インターバル	30
-	jpeg_scale_interval	JPEG 画像拡大率	1.1
		<*****ユーザが変更出来ないデータ*****>	
-	rule_path	利用規約チェック用パス。 規定値 : C:¥¥pdfstcorrector。	-
-	startup_path	多重起動防止用ファイルパス。 規定値 : C:¥¥pdfstcorrector¥¥multistartup.txt。	-

-	canvas_margin	シェルに対するキャンバス配置のマージン (X;Y)	60;60
-	slider_comp	スライダーコンポーネントの属性値	20;20;50;0; 100
-	dispatch	ディスパッチ回数指定(メッセージ用)	30;1
-	option_margin	オプション設定シェル(画面)に対するマージン	70
-	jpeg_margin	JPEG 変換ページ設定シェル(画面)に対するマージン	300;330;35
-	textupdate_canvas_margin	テキスト更新用シェル(画面)に対するキャンバス配置のマージン	675;105
-	textupdate_text_margin	テキスト更新用シェル(画面)に対するテキスト配置のマージン	355

プロパティファイルの更新

プロパティファイルは“Windows-31J”文字コードで読み書きしていますので、一般的な Windows 環境のテキストエディタで編集が可能です。

プロパティファイルの変更ルールは以下の通りです。

- 利用者が変更できる項目は限定されています。
- 変更可能かどうかは本ドキュメントやプロパティファイルコメントで明確化しています。
- 配布物内の初期プロパティファイルをそのまま使用しても通常は問題ありません。
(ただし PdfReader パス指定については適切に設定して下さい。)

プロパティファイルの具体的な編集方法は以下の通りです。

- 「キー=値」形式で値を指定します。
- キーの追加、名称変更や順序変更はしないで下さい。
- キーは1カラム目から指定します。
- 値として認識しない不要な空白文字は入力しないで下さい。
- 通常1行で1キーを指定しますが、複数行になる場合は最終文字をバックシュラッシュ
(通常¥で表示されます)にして継続します。
- 複数值の区切り文字は基本的にセミコロン(;)を使用しています。
- メッセージ等に用いる改行は¥nを使用しています。
- コメント行は1カラム目に#を指定します。
- 編集が完了したら編集ツールからファイルを解放(クローズ)して下さい。

エラーファイル

本ツールでのシステムエラーメッセージは C:\pdfstcorrector¥error¥error.txt に出力されます。

エラーが発生した場合はこのファイルも参照して下さい。

(エラーレベルだけでなく、警告レベルのメッセージも出力されています。)

Q&A

Q1 Cドライブ以外のドライブに配置したい

A1 jar ファイル内 path プロパティファイルを編集して下さい。

また起動バッチファイルや dfSlantTextCorrector プロパティファイルをチェックし変更必要箇所を特定し対処して下さい。

連絡先

本ツールに関する連絡は以下のメールアドレスへお願いします。

sqlcstool@yahoo.co.jp

ただし以下の注意点に留意して下さい。

- 個々のメールに返信することはありません。
- 誹謗中傷メールには着信拒否設定する場合があります。
- 操作方法に対する問い合わせには対応いたしません。

本利用者マニュアルを参照して下さい。

- 不具合情報に関しては確認・検証可能だった件についてのみ本ツールインターネットサイトに掲示され、対応が必要と本ツール開発者が判断したものについては修正・更新が行われます。

判断の根拠についての説明は行いません。

インターネットサイト URL <https://sqlcrspsite.web.fc2.com/>

確認・検証が不可能だった情報に関しては、とりあえず公開しておいた方が良いと判断したものについては利用者情報としてサイトに掲載されます。判断の根拠についての説明は行いません。

- 改善要望については本ツール開発者が必要と判断したものについて対応が行われます。

判断の根拠についての説明は行いません。